

This Page Is Inserted by IFW Operations  
and is not a part of the Official Record

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning documents *will not* correct images,  
please do not report the images to the  
Image Problems Mailbox.**

# PATENT ABSTRACTS OF JAPAN

(11)Publication number : 08-235143  
(43)Date of publication of application : 13.09.1996

(51)Int.Cl. G06F 15/177

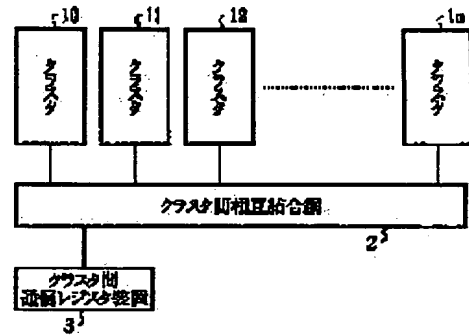
(21)Application number : 07-039545 (71)Applicant : KOFU NIPPON DENKI KK  
(22)Date of filing : 28.02.1995 (72)Inventor : ANDO NORIYUKI

## (54) PARALLEL COMPUTER OF CLUSTER CONSTITUTION

### (57)Abstract:

**PURPOSE:** To perform exclusive control, synchronous control, and communication control fast among arithmetic processors in different clusters.

**CONSTITUTION:** The parallel computer of cluster constitution composed of plural clusters 10-1n and an inter-cluster connecting network connecting them is provided with an inter-cluster communication register device 3 that the arithmetic processors in all the clusters can access for reading and writing, and this register device 3 is connected to the inter-cluster connecting network 2 to assign common variables used for the exclusive control, synchronous control, and communication control to the inter-cluster communication register device 3. Consequently, the time of access to the common variables is shortened and the access throughput is improved.



## LEGAL STATUS

[Date of request for examination] 28.02.1995  
[Date of sending the examiner's decision of rejection]  
[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]  
[Date of final disposal for application]  
[Patent number] 2731742  
[Date of registration] 19.12.1997  
[Number of appeal against examiner's decision of rejection]  
[Date of requesting appeal against examiner's decision of rejection]  
[Date of extinction of right]

(19)日本国特許庁 (J P)

(12) 公 開 特 許 公 報 (A)

(11)特許出願公開番号

特開平8-235143

(43)公開日 平成8年(1996)9月13日

(51)Int.Cl.<sup>6</sup>

G 0 6 F 15/177

識別記号

庁内整理番号

F I

G 0 6 F 15/16

技術表示箇所

4 0 0 S

審査請求 有 請求項の数4 O L (全 7 頁)

(21)出願番号 特願平7-39545

(22)出願日 平成7年(1995)2月28日

(71)出願人 000168285

甲府日本電気株式会社

山梨県甲府市大津町1088-3

(72)発明者 安藤 憲行

山梨県甲府市大津町1088-3 甲府日本電  
気株式会社内

(74)代理人 弁理士 京本 直樹 (外2名)

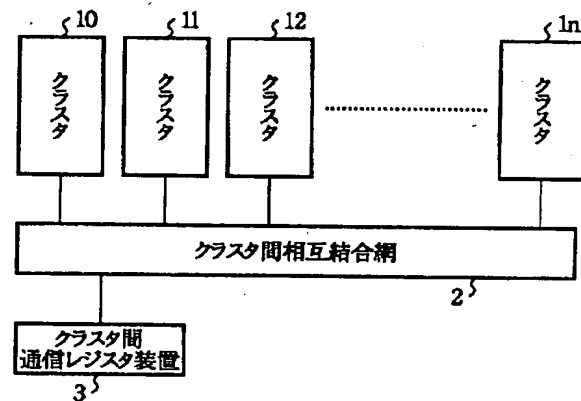
(54)【発明の名称】 クラスタ構成の並列計算機

(57)【要約】

【目的】 異なるクラスタの演算プロセッサ間の排他制御、同期制御、通信制御を高速に処理実行する。

【構成】 複数のクラスタ10~1nと、それらを結合するクラスタ間相互結合網2より構成されるクラスタ構成の並列計算機において、全てのクラスタ内の全ての演算プロセッサよりリードアクセス、およびライトアクセス可能なクラスタ間通信レジスタ装置3を設け、これをクラスタ間相互結合網2に接続することにより、排他制御、同期制御、通信制御で用いる共有変数を、このクラスタ間通信レジスタ装置3に割り当てる。

【効果】 共有変数のアクセスタイムの高速化、およびアクセススループットの向上ができる。



## 【特許請求の範囲】

【請求項1】 それぞれが、複数の演算プロセッサと、主記憶装置と、これら演算プロセッサ及び主記憶装置を接続するクラスタ内相互結合網とから成る複数のクラスタと、

該複数のクラスタを接続するクラスタ間相互結合網とで構成されるクラスタ構成の並列計算機において、

前記クラスタ間相互結合網に接続され、前記演算プロセッサから高速にリードアクセス及びライトアクセス可能な複数アドレスを有するクラスタ間通信レジスタ装置を設けて、専用のポートを付加し、

前記クラスタ間相互結合網は、前記クラスタからのリクエストがクラスタ間データ転送ならば行き先のクラスタ番号、またクラスタ間通信レジスタ装置アクセスならば前記クラスタ間通信レジスタのポート番号をそれぞれルーティングアドレスとする制御部を有し、

クラスタ間にわたる複数演算プロセッサ間の同期制御、排他制御及び通信制御には前記クラスタ間通信レジスタ装置を使用するよう構成された制御プログラムを実行することを特徴とするクラスタ構成の並列計算機。

【請求項2】 前記各クラスタ内において、前記クラスタ内相互結合網に接続される高速のクラスタ内通信レジスタ装置を付加し、

前記制御プログラムは、同一クラスタ内の演算プロセッサ間の同期制御、排他制御及び通信制御なら前記クラスタ内通信レジスタ装置を使用するよう最適化されたことを特徴とする請求項1記載のクラスタ構成の並列計算機。

【請求項3】 前記各クラスタ内において、前記クラスタ内相互結合網に接続され、前記クラスタ間通信レジスタ装置の写しを保持しておき、前記クラスタ間通信レジスタ装置に対するリードアクセス時に読み出される高速のクラスタ間通信レジスタコピー装置を付加したことを特徴とする請求項1及び請求項2記載のクラスタ構成の並列計算機。

【請求項4】 前記クラスタ間相互結合網又は前記クラスタ内相互結合網をクロスバースイッチ構成としたことを特徴とする請求項1又は請求項2記載のクラスタ構成の並列計算機。

## 【発明の詳細な説明】

## 【0001】

【産業上の利用分野】 本発明はクラスタ構成の並列計算機に関する。

## 【0002】

【従来の技術】 従来のクラスタ構成の並列計算機は、複数のクラスタと、それらを結合するクラスタ間相互結合網より構成されている。クラスタとは、複数の演算プロセッサ、主記憶装置およびクラスタ間相互結合網間を接続するクラスタ内相互結合網より構成される並列計算機ノードのことを言う。

【0003】 クラスタ間相互結合網は、結合網タイプとして、例えばフルクロスバ（スロスバースイッチ）を選択することができるが、この結合網はクラスタ間の通信（アクセス）のルーティングしか行わない。

【0004】 従って、異なるクラスタに属する複数の演算プロセッサ間において、排他制御、同期制御、通信制御を行うためには、例えば、富田眞治・末吉敏則共著「並列処理マシン」（電子情報通信学会編、オーム社発行）P. 117～120にも記載されている技術を使用して、いずれか一方のクラスタ内の主記憶装置に、一方の演算プロセッサが共有変数値を書き込み、これをもう一方の演算プロセッサが読み出すことが必要になる。つまり、2つの演算プロセッサの内、一方は自分のクラスタ内の主記憶装置へのアクセス（これをローカルアクセスと言う）となり、もう一方は他方のクラスタ内の主記憶装置へのアクセス（これをリモートアクセスと言う）になる。リモートアクセスの場合は、自プロセッサが属するクラスタとは異なるクラスタへアクセスする為、クラスタ間相互結合網を介してアクセス要求が伝達される。

【0005】 一般にクラスタ内の複数の演算プロセッサと主記憶装置を接続する結合網は、比較的速いアクセスタイムと高いメモリスループットが得られるように密に結合されるよう構成される。一方、複数のクラスタを接続する結合網は、実装技術および回路技術等の制約より、比較的遅いアクセスタイムと低い通信スループットで疎に結合される。従って、リモートアクセスはローカルアクセスに比べ、アクセス性能が低く、クラスタ構成の並列計算機の性能を向上する為には、リモートアクセスの頻度をなるべく低く抑えることがポイントとなる。

【0006】 また、クラスタ構成の並列計算機に限らず、複数の演算プロセッサを持つ並列計算機においては、並列処理を行おうとすると、異なる演算プロセッサ間で共有データのやり取りが発生する。特に、同期制御、排他制御、通信制御は、並列計算機で実行する並列処理において、並列実行が十分なされない所であり、高並列になるに従い、これら制御の全体に及ぼす影響は非常に大きくなる。従って、これら制御の処理時間を低減することは、並列計算機の性能向上に及ぼす効果は非常に大きいものである。

## 【0007】

【発明が解決しようとする課題】 上述した従来のクラスタ構成の並列計算機では、同期制御、排他制御、通信制御を行う場合、これらの制御で用いる共有変数の確保される領域は、クラスタ内の主記憶装置となる。この場合、他クラスタからのアクセスはリモートアクセスとして、クラスタ内演算プロセッサから同クラスタ内主記憶装置へのアクセスであるローカルアクセスとは別制御で処理される。ところで、リモートアクセスは一般的にア

クセスタイムは、クラスタ間の物理的距離の遠さによるアクセスディレーと、リモートアクセス先クラスタ内でのクラスタ内アクセスとの競合ディレーにより、非常に遅いものになる。従って、あるクラスタの主記憶に共有変数を割り当てた場合、頻繁にリモートアクセスが必要となる為、上記制御のオーバヘッドは非常に大きなものとなり、システム全体の性能を劣化させる要因となるという問題点がある。

【0008】

【課題を解決するための手段】第1の本発明は、それぞれ、複数の演算プロセッサと、主記憶装置と、これら演算プロセッサ及び主記憶装置を接続するクラスタ内相互結合網とから成る複数のクラスタと、該複数のクラスタを接続するクラスタ間相互結合網とで構成されるクラスタ構成の並列計算機において、前記クラスタ間相互結合網に接続され、前記演算プロセッサから高速にリードアクセス及びライトアクセス可能な複数アドレスを有するクラスタ間通信レジスタ装置を設けて、専用のポートを付加し、前記クラスタ間相互結合網は、前記クラスタからのリクエストがクラスタ間データ転送ならば行き先のクラスタ番号、またクラスタ間通信レジスタ装置アクセスならば前記クラスタ間通信レジスタのポート番号をそれぞれルーティングアドレスとする制御部を有し、クラスタ間にわたる複数演算プロセッサ間の同期制御、排他制御及び通信制御には前記クラスタ間通信レジスタ装置を使用するよう構成された制御プログラムを実行することを特徴とする。

【0009】第2の本発明は、第1の本発明に対して、各クラスタ内において、前記クラスタ内相互結合網に接続される高速のクラスタ内通信レジスタ装置を付加し、前記制御プログラムは、同一クラスタ内の演算プロセッサ間の同期制御、排他制御及び通信制御なら前記クラスタ内通信レジスタ装置を使用するように最適化されたことを特徴とする。

【0010】第3の本発明は、第1の本発明又は第2の本発明に対して、各クラスタ内において、前記クラスタ内相互結合網に接続され、前記クラスタ間通信レジスタ装置の写しを保持しておき、前記クラスタ間通信レジスタ装置に対するリードアクセス時に読み出される、高速のクラスタ間通信レジスタコピー装置を付加したことを特徴とする。

【0011】第4の本発明は、第1の本発明におけるクラスタ間相互結合網又は第2の本発明のクラスタ内相互結合網をクロスバスイッチ構成としたことを特徴とする。

【0012】

【作用】本発明は、主記憶装置より高速なクラスタ間通信レジスタ装置を設けてクラスタ間相互結合網に接続して、同期制御、排他制御、通信制御等を行う場合の共有変数の確保領域とすることにより、共有変数のアクセス

タイムの高速化とアクセススループットの向上が可能化する。

【0013】

【実施例】以下、本発明の実施例を図面を用いて詳細に説明する。

【0014】図1は第1の本発明のクラスタ構成の並列計算機のシステム構成例を示している。10、11、…、1nは各クラスタを示す。クラスタ10～1nは単体だけで、図2に示すように1つの並列計算機を構成している。2は各クラスタを結合するクラスタ間相互結合網を示している。各クラスタ10～1nはクラスタ間相互結合網2に対し1ポート分用意され、このポートとクラスタ間結合網2はリンクで結合される。従って、クラスタ間相互結合網はクラスタの個数をnとするならば、n入力、n出力のネットワーク構成になる。3は本発明で設けられたクラスタ間通信レジスタ装置を示す。

【0015】クラスタ間のリクエストの通信は全て、クラスタ間相互結合網2を経由することになる。クラスタ間相互結合網2は、同一のクラスタヘリクエスト要求をする複数個のリクエストが同時に到着した場合に対し、これを競合調停し、リクエストのルーティングを行う機能を有している。

【0016】図2はクラスタ10～1nそれぞれを構成例を示している。200、201、…、20mは演算プロセッサを示しているの、n個のクラスタで各クラスタにm子の演算プロセッサを有するならば、並列計算機システムでn×m個の演算プロセッサを持っていることになる。230はクラスタ内の主記憶装置を示す。この主記憶装置230をアクセスすることができるのは、自クラスタ内の演算プロセッサと、他クラスタの演算プロセッサ両方が可能であるとする。自クラスタ内の演算プロセッサが自クラスタ内の主記憶装置をアクセスすることをローカルアクセス、他クラスタの主記憶装置をアクセスすることをリモートアクセスと言う。

【0017】220はクラスタ内相互結合網を示す。クラスタ内相互結合網220は自クラスタ内の各演算プロセッサ210～21mと主記憶装置230を結合する。また、1ポート分がクラスタ間相互結合網2へのアクセスパスとして用意されている。従って、クラスタ間相互結合網2は、自クラスタ内の複数個の演算プロセッサから送られるローカルアクセスと、クラスタ間相互結合網より送られるリモートアクセスに対して、リクエストの競合調停を行い、各リクエストを所望の主記憶装置、演算プロセッサ、クラスタ間相互結合網2にルーティングする機能を有している。

【0018】他クラスタの主記憶装置230ヘリモートアクセスを行う場合には、リクエストの経路は、自演算プロセッサを起点として、自クラスタ内のクラスタ内相互結合網220、クラスタ間相互結合網2、他クラスタのクラスタ内相互結合網220を経て、他クラスタの主

記憶装置230へ至り、アクセス後のリブライは、上記バスの逆向きにアクセスが經由されることになる。

【0019】図3はクラスタ間通信レジスタ装置3の構成例を示す。このクラスタ間通信レジスタ装置3は、複数のワードより構成される通信レジスタとしてのRAM301と、通信レジスタアクセスを制御する通信レジスタ制御部310と、クラスタ間相互結合網2へのインターフェース機能をもつリクエストパケット制御部311及びリブライパケット制御部312を主な構成要素とする。RAM301は、主記憶装置230より高速で、0番地から連続的にアドレス番号が振られている。演算プロセッサからの通信レジスタアクセスでは、RAM301のアドレスを指定することにより、アクセスする通信レジスタのワードを決めることが出来る。

【0020】また、このRAM301をアクセスする為にライトレジスタ302、アドレスレジスタ303及びリードレジスタ304が備えられており、ライトイネーブル305及びリードイネーブル306が供給される。RAM301にアクセスしない場合は、ライトイネーブル、リードイネーブル共“0”値にする。

【0021】RAM301に対してライトを行う場合、ライトイネーブルを“1”値にし、ライトをするワードのアドレスをアドレスレジスタ303に入れ、ライトをするデータをライトレジスタ302に入れる。次にタイミングで、ライトレジスタ302の値はアドレスレジスタ303で指定されたワードに対しライトされる。

【0022】また、RAM301に対してリードを行う場合、リードイネーブルを“1”値にし、リードするワードのアドレスをアドレスレジスタ303に入れる。次のタイミングでアドレスレジスタ303で指定されたワードの値は、リードレジスタ304へ入れられる。これら、RAM周辺部の各レジスタの制御は、通信レジスタ制御部310で制御する。

【0023】リクエストパケット制御部311は、クラスタ間相互結合網2により到着するリクエストに対する処理を行う。相互結合網2よりリクエストを受け取ったならば、コードフィールドでデコードし、ロードアクセスなのかストアアクセスなのかも判断する。このデコード結果は通信レジスタ制御部310に送られる。リブライパケット制御部312は、RAM301に対し、リードレジスタ304内のデータをリブライのデータフィールドとして構成し、クラスタ間相互結合網2に対するリブライリクエストを生成する機能を有する。

【0024】次に、各通信レジスタアクセスにおける、クラスタ間通信レジスタ装置3内での処理方法について述べる。ストアアクセスならば、アドレスフィールドで示されたアドレスのRAM301のワードに対し、データフィールド内のデータを書き込む制御となる。即ち、書き込むタイミングにおいて、アドレスフィールド内の通信レジスタアドレスをアドレスレジスタ303に入れ

る。また、データフィールド内の書き込みデータをライトレジスタ302に入れる。同時にライトイネーブルを“1”値にすることにより、次にタイミングでストアアクセスが完了する。一方、ロードアクセスならば、アドレスフィールドで示されたアドレスのRAM301のワードの値を読みだし、これを、リブライのデータとし、クラスタ間相互結合網2に送出する制御となる。即ち、読み出すタイミングにおいて、アドレスフィールド内の通信レジスタアドレスをアドレスレジスタ303に入れ、同時にリードイネーブルを“1”値にする。次のタイミングで読み出されたデータがリードレジスタ308に書き込まれる。このリードレジスタ値は次のタイミングでリブライリクエスト制御部312に送られ、ここでリブライリクエストを構成し、クラスタ間相互結合網2に送出する。

【0025】図4は、クラスタ間相互結合網2の構成例を示す。本例ではクラスタ間相互結合網3の基本構成としてフルクロスバースイッチを用いている。クラスタ間相互結合網3の構成としては、バス、リング等多くのネットワーク構成が本発明の適応できるが、演算プロセッサAから通信レジスタへのリクエストと、演算プロセッサBからの通信レジスタ乙へのリクエストが同時に到達した時に、競合（ブロッキング）が発生しないネットワーク構成が望ましい。フルクロスバースイッチは全くブロッキングが生じない。

【0026】本フルクロスバースイッチは、 $(n+1)$ ポート入力 $(n+1)$ ポート出力であり、各々 $n$ 個のクラスタと、1個のクラスタ間通信レジスタ装置3に接続される。即ち、入力ポートの0から $n-1$ ポートは各クラスタに接続され、 $n$ ポートは通信レジスタ装置に接続される。出力ポートの接続方法も同様である。

【0027】400、400、…40 $n$ は $(n+1)$ 個の入力ポート、410、411、…41 $n$ は入力バッファを示している。420、421、…42 $n$ は入力ポートから来るリクエストと入力バッファからのリクエストを選択する2ウェイのセレクトである。430、431、…43 $n$ は $(n+1)$ ウェイのセレクト、440、441、…44 $n$ は $(n+1)$ 個の出力ポートを示している。450はクロスバースイッチ制御部を示している。

【0028】次に本クロスバースイッチでのリクエストルーティング処理について述べる。入力ポートよりクラスタが送出したリクエストが到着したならば、リクエスト内のルーティングアドレスフィールドをクロスバースイッチ制御部550に送る。このフィールドには $(n+1)$ 個の出力ポートのうち、どの出力ポートに通過したいかが書かれている。ここで、クラスタ間データ転送ならば、行き先のクラスタ番号がルーティングアドレスになり、クラスタ間通信レジスタアクセスならば、ルーティングアドレスは $n$ ポートとなっている。

【0029】クロスバースイッチ制御部450は同一の出力ポー

トに対し、同時に通過要求を出している複数のリクエストに対し、ポートの優先度に従い1つのリクエストを選択する。選択した入力ポート番号は、その出力ポート対応の(n+1)ウエイセクタのセレクト信号として送られ、選択されたリクエストは(n+1)ウエイセクタを通過し、出力ポートに達する。選択されなかったリクエストは、入力バッファに一時的に格納され、次のタイミングで再度クロスバ制御部450に対しリクエスト通過要求を出す。クロスバ制御部550は、故意にあるポートのリクエスト通過が遅れらせることがない10ように、公平なポート優先度を持つ必要がある。

【0030】図5は、第2の本発明のクラスタの構成例を示している。システム構成は図1に示した第1の本発明の構成と同一である。但し、本発明の各クラスタ内において、図5に示すように、クラスタ内通信レジスタ装置530が設けられる。クラスタ内相互結合網510は、クラスタ内通信レジスタ装置用にアクセスポートを1つ用意し、このポートとクラスタ内通信レジスタ装置530はアクセスバスで結合される。各演算プロセッサ500、501、…50mが発行するクラスタ内通信レジスタアクセスは、クラスタ内相互結合網510によりルーティングされ、上記アクセスバスを経由してクラスタ内通信レジスタ装置530に到達する。

【0031】但し、クラスタ内通信レジスタ装置530をアクセスできるのは、そのクラスタ内の演算プロセッサに限るものとする。即ち、他クラスタの演算プロセッサからのリモートアクセスを処理実行する機能は、クラスタ内相互結合網510およびクラスタ内通信レジスタ装置530には有していない。

【0032】クラスタ内通信レジスタ装置530の構成は、図3に示したクラスタ間通信レジスタ装置3と同一の構成である。即ち、クラスタ内通信レジスタ装置530も複数のワードより構成されるRAMと、通信レジスタアクセスを制御する通信レジスタ制御部と、クラスタ内相互結合網510へのインターフェース機能をもつリクエストパケット制御部及びリプライパケット制御部より構成される。RAMは0番地から連続的にアドレス番号が振られている。演算プロセッサからの通信レジスタアクセスでは、この通信レジスタアドレスを指定することにより、アクセスする通信レジスタのワードを決めることが出来る。

【0033】次に、第3の本発明として、クラスタ内に、図5に示したクラスタ内通信レジスタ装置の他に、図1に示したクラスタ間通信レジスタ装置のコピーを保持するクラスタ間通信レジスタコピー装置を設けたものが考えられる。

【0034】クラスタ内通信レジスタは第2の本発明におけるクラスタ内演算プロセッサ専用の通信レジスタであり、クラスタ間通信レジスタコピー装置はクラスタ間通信レジスタ装置と同一容量のワードを有するRAMに

より構成され、常にクラスタ間通信レジスタ装置と同一内容がコピーされている。

【0035】クラスタ間通信の場合、アクセスは次のように処理される。リードアクセスならばクラスタ間通信レジスタコピー装置に対しアクセスを行う。

【0036】一方、ライトアクセスならばクラスタ間通信レジスタ装置に対しアクセスを行う。テスト&セット系のアクセスもライト動作が入るのでライトアクセスと同じ処理となる。クラスタ間通信レジスタ装置はライト動作が入ると、ライトをしたRAMのアドレスとライトデータを制御リクエストとして構成し、これを全クラスタのクラスタ間通信レジスタコピー装置にブロードキャストする。各クラスタ間通信レジスタコピー装置はこの制御リクエストを受け取ったならば、そのアドレスが示す通信レジスタコピーの同一アドレスに対しデータ内容を書き込む。以上の処理を行うことで、クラスタ間通信レジスタ装置とクラスタ間通信レジスタコピー装置との一貫性を保証することができる。

【0037】

【発明の効果】第1の本発明は、以上に説明したように、共有変数割り当て領域を、クラスタの主記憶装置からクラスタ間通信レジスタへ変えることにより、主記憶装置へのリモートアクセスからクラスタ間通信レジスタアクセスにすることができる。従って、クラスタ間の物理的遠距離より生じるアクセスディレー分や、他クラスタ内競合によるディレー分を無くすることが可能となり、クラスタ間に渡る複数の演算プロセッサ間の同期制御、排他制御、通信制御等における共有変数アクセスを高速に実行可能となる。

【0038】また、第2の本発明は、クラスタ内通信レジスタとクラスタ間通信レジスタの2階層の通信レジスタを持つことになり、複数の演算プロセッサ間の同期、排他、通信制御において、それら制御プログラムの最適化を行うことで、より高速な実行が可能になる。例えば、同期制御、排他制御、通信制御を行うプログラムにおいて、これを実行する複数の演算プロセッサが同一クラスタに閉じているならば、そのクラスタ内のクラスタ内通信レジスタを使用するように変数を割り当てる。複数の演算プロセッサが異なるクラスタに渡ってしまう場合には、クラスタ間通信レジスタを使用するように変数を割り当てる。

【0039】また、システム的全演算プロセッサの同期制御を行う場合でも、通信レジスタの階層構造を用いることにより、高速な実行が可能となるプログラムの最適化ができる。これは、各クラスタにおいて、クラスタ内通信レジスタを用いて、クラスタ内全演算プロセッサの同期を取ってから、クラスタ代表の演算プロセッサがクラスタ間通信レジスタを用いて、クラスタの同期を取るようなプログラム構成にする。このようにすることにより、クラスタ内ではクラスタ内通信レジスタを用いた、

局所的に高速な同期処理が可能となり、また、クラスタ代表のみがクラスタ間通信レジスタにアクセスするので、アクセス競合が低減できる。

【0040】また、第3の本発明は、クラスタ内にクラスタ間通信レジスタのコピーをとっておくことにより、リードアクセスがクラスタ内の通信レジスタアクセスとなる為、スピンロック等で発生する頻発するリードアクセスの集中が各クラスタ毎に分散緩和され、同期制御、排他制御のオーバーヘッド削減に大きな効果を生むことができる。

【図面の簡単な説明】

【図1】第1の本発明の並列計算機システムの構成例を示す図である。

【図2】第1の本発明におけるクラスタの構成例を示す図である。

【図3】第1、第2、第3の本発明におけるクラスタ間通信レジスタ装置、クラスタ内通信レジスタ装置及びクラスタ間通信レジスタコピー装置の構成例を示す図である。

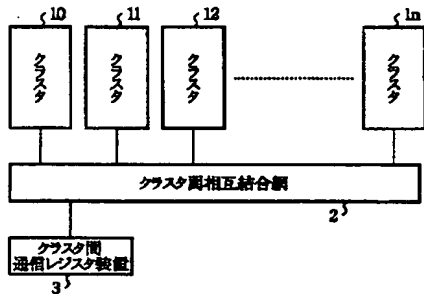
【図4】第1の本発明におけるクラスタ相互結合毛の構成例を示す図である。

【図5】第2の本発明におけるクラスタの構成例を示す図である。

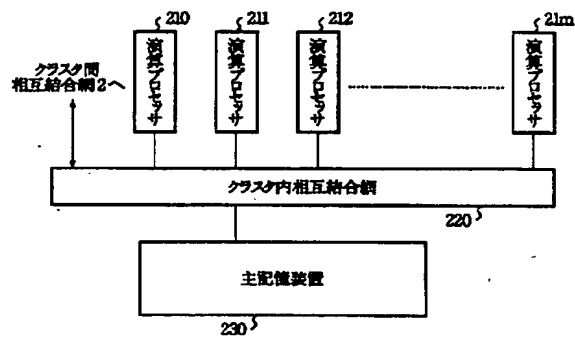
【符号の説明】

- 10, 11, 12~1n クラスタ
- 2 クラスタ間相互結合網
- 3 クラスタ間通信レジスタ装置
- 210, 211, 212~21n, 500, 501, 502~50n 演算プロセッサ
- 220, 510 クラスタ内通信レジスタ装置
- 230, 520 主記憶装置
- 301 RAM
- 302 ライトレジスタ
- 303 アドレスレジスタ
- 304 リードレジスタ
- 305 リードイネーブル
- 306 ライトイネーブル
- 310 通信レジスタ制御部
- 311 リクエストパケット制御部
- 312 リプライパケット制御部
- 400, 401~40n 入力ポート
- 410, 411~41n 入力バッファ
- 420, 421~42n 2ウェイセクタ
- 430, 431~43n (n+1) ウェイセクタ
- 440, 441~44n 出力ポート
- 450 クロスバー制御部。

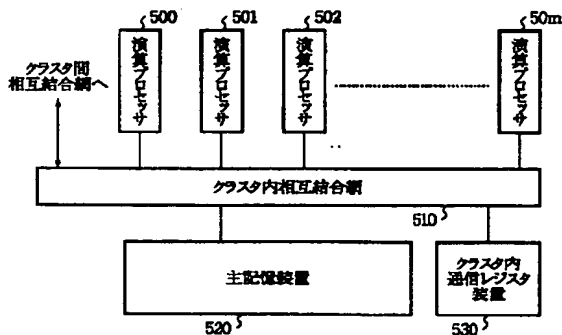
【図1】



【図2】

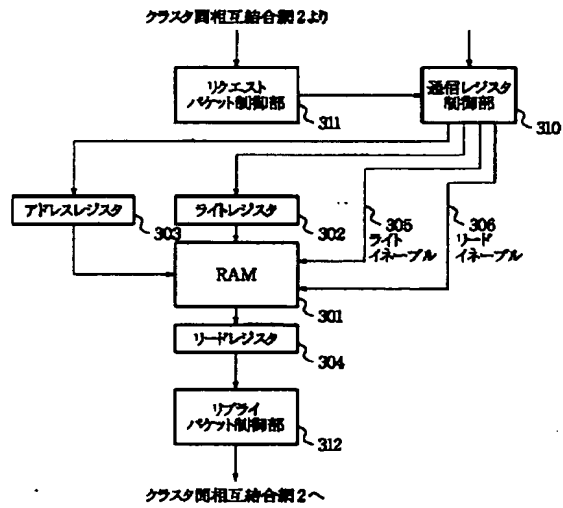


【図5】





【図3】



【図4】

